WHAT IS CLAIMED IS:

1. A data processing apparatus for providing a browser apparatus with the contents of data provided on a network in a form of voice data, comprising:

means for forming, on the basis of the data provided on said network, voice data indicating a part or the whole of the contents of the data;

means for storing the formed voice data;

means for forming data by adding to the data provided on said network an identifier indicating a location where the voice data is stored; and

means for providing said browser apparatus with the data to which the identifier is added.

2. A data processing apparatus for permitting a browser apparatus to respond by voice to data provided on a network, comprising:

means for checking whether the contents of the data provided on said network include a content requiring a response from said browser apparatus;

means for forming data by adding to the data provided on said network an identifier indicating a recipient of the response sent by voice data from said browser apparatus; and

means for providing said browser apparatus with the data to which the identifier is added.

3. The apparatus according to claim 2, further comprising recognizing means for performing voice

recognition for voice data related to the response, when the voice data is supplied from said browser apparatus to said recipient.

4. The apparatus according to claim 3, further comprising:

means for forming response data in a form suited to a server for receiving the response on said network, on the basis of the result of recognition by said recognizing means; and

means for providing the response data to said server.

5. The apparatus according to claim 2, further comprising:

means for forming a recognition grammar for recognizing voice data related to each of a plurality of predetermined items, when the response is to be selected from said plurality of items;

means for determining, on the basis of the recognition grammar, to which item the voice data related to the response from said browser apparatus corresponds;

means for forming response data in a form suited to a server for receiving the response on said network, in accordance with each item; and

means for providing the response data to said server.

6. The apparatus according to claim 5, wherein the

response data is formed before data to which the identifier is added is provided to said browser apparatus.

7. A browser system comprising a browser apparatus, a server for providing data to said browser apparatus via a network, and a data processing apparatus for providing said browser apparatus with the contents of data provided by said server in a form of voice data, wherein

said data processing apparatus comprises:

means for forming, on the basis of the data provided by said server, voice data indicating a part or the whole of the contents of the data;

means for storing the formed voice data;

means for forming data by adding to the data provided by said server an identifier indicating a location where the voice data is stored; and

means for providing said browser apparatus with the data to which the identifier is added, and

said browser apparatus comprises means for acquiring the voice data from the location indicated by the identifier and outputting a voice related to the voice data.

8. A browser system comprising a browser apparatus, a server for providing data to said browser apparatus via a network, and a data processing apparatus for permitting the browser apparatus to respond by voice to

data provided by said server, wherein

said data processing apparatus comprises:

means for checking whether the contents of the data provided on said network include a content requiring a response from said browser apparatus;

means for forming data by adding to the data provided by said server an identifier indicating a recipient of the response sent by voice data from said browser apparatus;

means for providing said browser apparatus with the data to which the identifier is added;

recognizing means for performing voice recognition for voice data related to the response, when the voice data is supplied from said browser apparatus to said recipient;

means for forming response data in a form suited to said server for receiving the response, on the basis of the result of recognition by said recognizing means; and

means for providing the response data to said server, and

said browser apparatus comprises:

means for inputting a voice;

means for forming voice data on the basis of the input voice; and

means for supplying the formed voice data to a recipient indicated by the identifier.

9. A browser system comprising a browser apparatus, a server for providing data to said browser apparatus via a network, and a data processing apparatus for providing the contents of data provided by said server in a form of voice data to said browser apparatus, and permitting said browser apparatus to respond by voice to data provided by said server, wherein

said data processing apparatus comprises:

means for forming, on the basis of the data provided by said server, voice data indicating a part or the whole of the contents of the data;

means for storing the formed voice data;

means for forming data by adding to the data provided by said server a first identifier indicating a location where the voice data is stored;

means for providing said browser apparatus with the data to which the first identifier is added;

means for checking whether the contents of the data provided by said server include a content requiring a response from said browser apparatus;

means for forming data by adding to the data provided by said server a second identifier indicating a recipient of the response sent by voice data from said browser apparatus;

means for providing said browser apparatus with the data to which the identifier is added;

recognizing means for performing voice

recognition for voice data related to the response, when the voice data is supplied from said browser apparatus to said recipient;

means for forming response data in a form suited to said server for receiving the response, on the basis of the result of recognition by said recognizing means; and

means for providing the response data to said server, and

said browser apparatus comprises:

means for acquiring the voice data from the location indicated by the first identifier and outputting a voice related to the voice data;

means for inputting a voice;

means for forming voice data on the basis of the input voice; and

means for supplying the formed voice data to a recipient indicated by the second identifier.

10. A data processing method of providing a browser apparatus with the contents of data provided on a network in a form of voice data, comprising the steps of:

forming, on the basis of the data provided on the network, voice data indicating a part or the whole of the contents of the data;

storing the formed voice data;

forming data by adding to the data provided on

the network an identifier indicating a location where the voice data is stored; and

providing the browser apparatus with the data to which the identifier is added.

11. A data processing method of permitting a browser apparatus to respond by voice to data provided on a network, comprising the steps of:

checking whether the contents of the data provided on the network include a content requiring a response from the browser apparatus;

forming data by adding to the data provided on the network an identifier indicating a recipient of the response sent by voice data from the browser apparatus; and

providing the browser apparatus with the data to which the identifier is added.

12. The method according to claim 11, further comprising the recognition step of performing voice recognition for voice data related to the response, when the voice data is supplied from the browser apparatus to the recipient.

13. The method according to claim 12, further comprising the steps of:

forming response data in a form suited to a server for receiving the response on the network, on the basis of the result of recognition in the recognition step; and

providing the response data to the server.

14. The method according to claim 11, further comprising the steps of:

forming a recognition grammar for recognizing voice data related to each of a plurality of predetermined items, when the response is to be selected from the plurality of items;

determining, on the basis of the recognition grammar, to which item the voice data related to the response from the browser apparatus corresponds;

forming response data in a form suited to a server for receiving the response on the network, in accordance with each item; and

providing the response data to the server.

15. The method according to claim 14, wherein the response data is formed before data to which the identifier is added is provided to the browser apparatus.

16. A recording medium recording a program which, in order to provide a browser apparatus with the contents of data provided on a network in a form of voice data, allows a computer to function as:

means for forming, on the basis of the data provided on said network, voice data indicating a part or the whole of the contents of the data;

means for storing the formed voice data;

means for forming data by adding to the data

provided on said network an identifier indicating a location where the voice data is stored; and

means for providing said browser apparatus with the data to which the identifier is added.

17. A recording medium recording a program which, in order to permit a browser apparatus to respond by voice to data provided on a network, allows a computer to function as:

means for checking whether the contents of the data provided on said network have contents requiring a response from said browser apparatus;

means for forming data by adding to the data provided on said network an identifier indicating a recipient of the response sent by voice data from said browser apparatus; and

means for providing said browser apparatus with the data to which the identifier is added.

18. The medium according to claim 17, wherein said program comprises a program which allows a computer to function as recognizing means for performing voice recognition for voice data related to the response, when the voice data is supplied from said browser apparatus to said recipient.

19. The medium according to claim 18, wherein said program comprises a program which allows a computer to function as:

means for forming response data in a form suited

to a server for receiving the response on said network, on the basis of the result of recognition by said recognizing means; and

means for providing the response data to said server.

20. The medium according to claim 17, wherein said program comprises a program which allows a computer to function as:

means for forming a recognition grammar for recognizing voice data related to each of a plurality of predetermined items, when the response is to be selected from said plurality of items;

means for determining, on the basis of the recognition grammar, to which item the voice data related to the response from said browser apparatus corresponds;

means for forming response data in a form suited to a server for receiving the response on said network, in accordance with each item; and

means for providing the response data to said server.

21. The medium according to claim 20, wherein the response data is formed before data to which the identifier is added is provided to said browser apparatus.

22. The apparatus according to claim 1, wherein the data provided on said network is described in a markup

language, and the identifier is added to the data as a tag corresponding to the markup language.

23. The apparatus according to claim 2, wherein the data provided on said network is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

24. The system according to claim 7, wherein the data provided by said server is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

25. The system according to claim 8, wherein the data provided by said server is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

26. The system according to claim 9, wherein the data provided by said server is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

27. The method according to claim 10, wherein the data provided on said network is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

28. The method according to claim 11, wherein the data provided on said network is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

29. The medium according to claim 16, wherein the

data provided on said network is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

30. The medium according to claim 17, wherein the data provided on said network is described in a markup language, and the identifier is added to the data as a tag corresponding to the markup language.

31. A browser apparatus comprising:

means for inputting a voice;

means for forming voice data on the basis of the input voice; and

means for supplying the formed voice data to a recipient indicated by a given identifier.

32. The apparatus according to claim 26, further comprising means for acquiring voice data from a location indicated by a given second identifier, and outputting a voice related to the voice data.

33. A data processing apparatus capable of communicating with a server and a browser apparatus via a network, comprising:

means for forming, on the basis of data provided by said server, voice data indicating a part or the whole of the contents of the data;

means for storing the formed voice data;

means for adding to the data provided by said server a first identifier indicating a location where the voice data is stored;

means for checking whether the contents of the data provided by said server include a content requiring a response from said browser apparatus;

means for further adding, when the contents of the data provided by said server have contents requiring a response, a second identifier indicating a recipient of the response to the data to which the first identifier is added;

means for providing said browser apparatus with the data to which the first identifier or the first and second identifiers are added;

recognizing means for performing voice recognition for voice data related to the response, when the voice data is supplied from said browser apparatus to said recipient;

means for forming response data in a form suited to said server for receiving the response, on the basis of the recognition result by said recognizing means; and

means for providing the response data to said server.